Identification of Genes Encoding the Folate- and Thiamine-Binding Membrane Proteins in Firmicutes

Eudes, Aymerick; Erkens, Guus B.; Slotboom, Dirk J.; Rodionov, Dmitry A.; Naponelli, Valeria; Hanson, Andrew D.

Published in:
Journal of Bacteriology

DOI:
10.1128/JB.01070-08

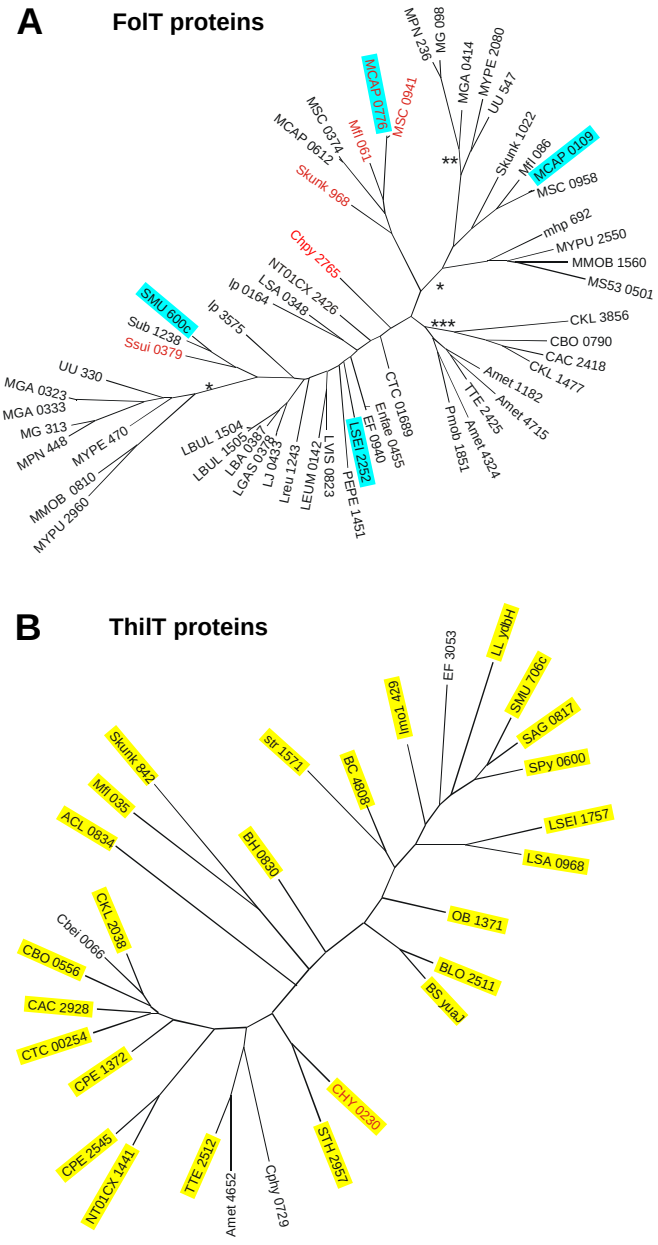
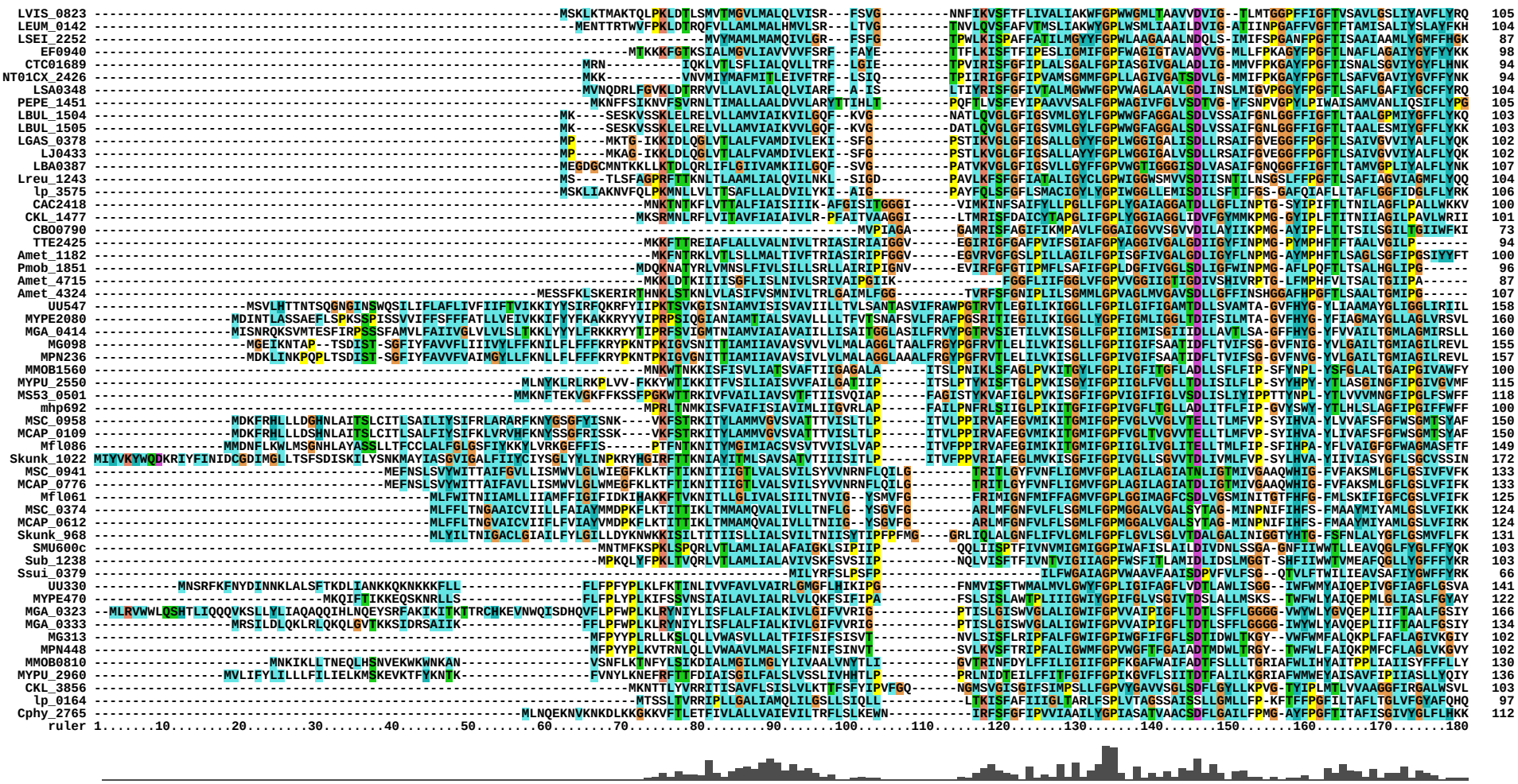
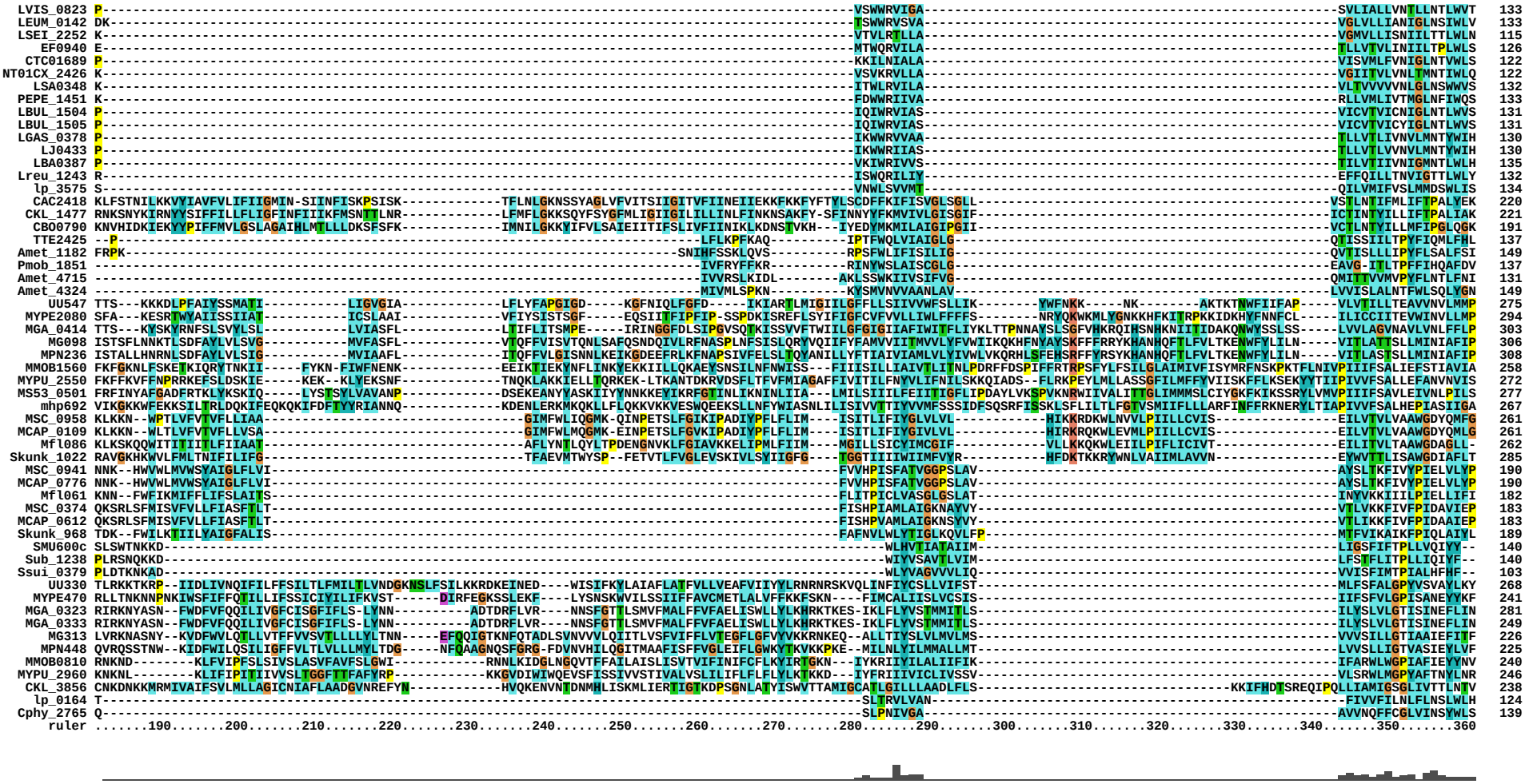


FIG. S1. Phylogenetic analysis of FolT and ThiT family proteins. After sequence alignment with ClustalW, the PHYLIP package was used to construct maximum likelihood trees, including bootstrapping with 100 replicates and drawing of a consensus tree. Genome abbreviations are listed in Supplemental Table 1. (A) Tree of FolT family proteins. The *folT* genes encoding proteins shown in red cluster with the folate salvage gene *folC*. Proteins highlighted in blue were shown in this study to bind folates. Asterisks indicate special features, as follows: *, two groups of *Mycoplasma* proteins with an insertion (~100-residues) containing two transmembrane domains after the third transmembrane domain; **, three *Mycoplasma* proteins with an additional hydrophilic C-terminal domain (~100 residues) similar to aspartyltRNA amidotransferase subunit C; ***, four *Clostridium* proteins with an insertion (~85 residues) containing two transmembrane domains after the third transmembrane domain. (B) Tree of ThiT family proteins. The *thiT* gene encoding the protein shown in red clusters with the thiamine salvage gene *thiN* (thiamine pyrophosphokinase). Proteins highlighted in yellow are encoded by *thiT* genes that are preceded by a thiamine pyrophosphate riboswitch.

Fig. S2. Sequence alignment of FolT proteins

Histograms below alignment show the degree of residue conservation at each position




LVIS_0823 P VSWWRVIGA SVLIALLVNTLLNTLWVT 133
LEUM_0142 DK TSWWRVSVA VGLVLLIANIGLNSIWLV 133
LSEI_2252 K VTVLRTLLA VGMVLLISNIILTTLWLN 115
EF0940 E MTWQRVILA TLLVTVLINIILTPLWLS 126
CTC01689 P KKILNIALA VISVMLFVNIGLNTVWLS 122
NT01CX_2426 K VSVKRVLLA VGIITVLVNLTMNTIWLQ 122
LSA0348 K ITWLRVILA VLTVVVVVNLGLNSWWVS 132
PEPE_1451 K FDWWRIIVA RLLVMLIVTMGLNFIWQS 133
LBUL_1504 P IQIWRVIAS VICVTVICNIGLNTLWVS 131
LBUL_1505 P IQIWRVIAS VICVTVICYIGLNTLWVS 131
LGAS_0378 P IKWWRVVAA TLLVTLIVNVLMNTYWIH 130
LJ0433 P IKWWRIIAS TLLVTLVVNVLMNTYWIH 130
LBA0387 P VKIWRIVVS TILVTIIVNIGMNTLWLH 135
Lreu_1243 R ISWQRILIY EFFQILLTNVIGTTLWLY 132
lp_3575 S VNWLSVVMI QILVMIFVSLMMDSWLIS 134
CAC2418 KLFSTNILKKVYIAVFVLIFIIGMIN-SIINFISKPSISK----------TFLNLGKNSSYAGLVFVITSIIGITVFIINEIIEKKFKKFYFTYLSCDFFKIFISVGLSGLL VSTLNTIFMLIFTPALYEK 220
CKL_1477 RNKSNYKIRNYYSIFFILLFLIGFINFIIIKFMSNTTLNR----------LFMFLGKKSQYFSYGFMLIGIIGILILLINLFINKNSAKFY-SFINNYYFKMVIVLGISGIF ICTINTYILLIFTPALIAK 221
CBO0790 KNVHIDKIEKYYPIFFMVLGSLAGAIHLMTLLLDKSFSFK----------IMNILGKKYIFVLSAIEIITIFSLIVFIINIKLKDNSTVKH---IYEDYMKMILAIGIPGII VCTLNTYILLMFIPGLQGK 191
TTE2425 --P LFLKPFKAQ---------IPTFWQLVIAIGLG QTISSIILTPYFIQMLFHL 137
Amet_1182 FRPK SNIHFSSKLQVS---------RPSFWLIFISILIG QVTISLLLIPYFLSALFSI 149
Pmob_1851 IVFRYFFKR---------RINYWSLAISCGLG EAVG-ITLTPFFIHQAFDV 137
Amet_4715 IVVRSLKIDL---------AKLSSWKIIVSIFVG QMITTVVMVPYFLNTLFNI 131
Amet_4324 MIVMLSPKN---------KYSMVNVVAANLAV LVVISLALNTFWLSQLYGN 149
UU547 TTS---KKKDLPFAIYSSMATI----------LIGVGIA----------LFLYFAPGIGD-----KGFNIQLFGFD-----IKIARTLMIGIILGFFLLSIIVVWFSLLIK--------YWFNKK-----NK-------AKTKTNWFIIFAP-----VLVTILLTEAVVNVLMMP 275
MYPE2080 SFA---KESRTWYAIISSIIAT----------ICSLAAI----------VFIYSISTSGF-----EQSIITFIPFIP-SSPDKISREFLSYIFIGFCVFVVLLIWLFFFFS-------NRYQKWKMLYGNKKHFKITRPKKIDKHYFNNFCL-----ILICCIITEVWINVLLMP 294
MGA_0414 TTS---KYSKYRNFSLSVYLSL----------LVIASFL----------LTIFLITSMPE-----IRINGGFDLSIPGVSQTKISSVVFTWIILGFGIGIIAFIWITFLIYKLTTPNNAYSLSGFVHKRQIHSNHKNIITIDAKQNWYSSLSS-----LVVLAGVNAVLVNLFFLP 303
MG098 ISTSFLNNKTLSDFAYLVLSVG----------MVFASFL----------VTQFFVISVTQNLSAFQSNDQIVLRFNASPLNFSISLQRYVQIIFYFAMVVIITMVVLYFVWIIKQKHFNYAYSKFFRRYKHANHQFTLFVLTKENWFYLILN-----VITLATTSLLMINIAFIP 306
MPN236 ISTALLHNRNLSDFAYLVLSIG----------MVIAAFL----------ITQFFVLGISNNLKEIKGDEEFRLKFNAPSIVFELSLTQYANILLYFTIAIVIAMLVLYIVWLVKQRHLSFEHSRFFYRSYKHANHQFTLFVLTKENWFYLILN-----VITLASTSLLMINIAFIP 308
MMOB1560 FKFGKNLFSKETKIQRYTNKII-----FYKN-FIWFNENK----------EEIKTIEKYNFLINKYEKKIILLQKAEYSNSILNFNWISS---FIIISILLIAIVTLITNLPDRFFDSPIFFRTRPSFYLFSILGLAIMIVFISYMRFNSKPKTFLNIVPIIIFSALIEFSTIAVIA 258
MYPU_2550 FKFFKVFFNPRRKEFSLDSKIE-----KEK--KLYEKSNF----------TNQKLAKKIELLTQRKEK-LTKANTDKRVDSFLTFVFMIAGAFFIVITILFNYVLIFNILSKKQIADS--FLRKPEYLMLLASSGFILMFFYVIISKFFLKSEKYYTIIPIVVFSALLEFANVNVIS 272
MS53_0501 FRFINYAFGADFRTKLYKSKIQ-----LYSTSYLVAVANP----------DSEKEANYYASKIIYYNNKKEYIKRFGTINLIKNINLIIA---LMILSIIILFEIITIGFLIPDAYLVKSPVKNRWIIVALITTGLIMMMSLCIYGKFKIKSSRYLVMVPIIIFSAVLEIVNLPILS 277
mhp692 VIKGKKWFEKKSILTRLDQKIFEQKQKIFDFIYYRIANNQ----------KDENLERKMKQKLLFLQKKVKKVESWQEEKSLLNFYWIASNLILISIVVTTIYVVMFSSSIDFSQSRFISSKLSFLILTLFGTVSMIIFLLLARFINFFRKNERYLTIAPIVVFSALHEPIASIIGA 267
MSC_0958 KLKKN--WPTLVFVTVFLLIAA----------GIMFWLIQGMK-QINPETSLFGIKIPADIYPFLFLIM----ISITLIFIYGLVLVL----------HIKKRDKWLNVVLPIILLCVIS----------EILVTVLVAAWGDYQMFG 261
MCAP_0109 KLKKN--WLTLVFVTVFLLVSA----------GIMFWLMQGMK-EINPETSLFGVKIPADIYPFLFLIM----ISITLIFIYGIVLVL----------HIRKRQKWLEVMLPIILLCVIS----------EILVTVLVAAWGDYQMLG 261
Mfl086 KLKSKQQWITITIITLFIIAAT----------AFLYNTLQYLTPDENGNVKLFGIAVKKELIPMLFIIM----MGILLSICYIMCGIF----------VLLKKQKWLEIILPIFLICIVT----------EILITVLTAAWGDAGLL- 262
Skunk_1022 RAVGKHKWVLFMLTNIFILIFG----------TFAEVMTWYSP--FETVTLFVGLEVSKIVLSYIIGFG----TGGTIIIIWIIMFVYR----------HFDKTKKRYWNLVAIIMLAVVN----------EYWVTTLISAWGDIAFLT 285
MSC_0941 NNK--HWVWLMVWSYAIGLFLVI----------FVVHPISFATVGGPSLAV AYSLTKFIVYPIELVLYP 190
MCAP_0776 NNK--HWVWLMVWSYAIGLFLVI----------FVVHPISFATVGGPSLAV AYSLTKFIVYPIELVLYP 190
Mfl061 KNN--FWFIKMIFFLIFSLAITS----------FLITPICLVASGLGSLAT INYVKKIIILPIELLIFI 182
MSC_0374 QKSRLSFMISVFVLLFIASFTLT----------FISHPIAMLAIGKNAYVY VTLVKKFIVFPIDAVIEP 183
MCAP_0612 QKSRLSFMISVFVLLFIASFTLT----------FISHPVAMLAIGKNSYVY VTLIKKFIVFPIDAAIEP 183
Skunk_968 TDK--FWILKTIILYAIGFALIS----------FAFNVLWLYTIGLKQVLFP MTFVIKAIKFPIQLAIYL 189
SMU600c SLSWTNKKD----------WLHVTIATAIIM LIGSFIFTPLLVQIYY-- 140
Sub_1238 PLRSNQKKD----------WIYVSAVTLVIM LFSTFLITPLLIQIYF-- 140
Ssui_0379 PLDTKNKAD----------WLYVAGVVVLIQ VVISFIMTPIALHFHF-- 103
UU330 TLRKKTKRP--IIDLIVNQIFILFFSILTLFMILTLVNDGKNSLFSILKKRDKEINED----WISIFKYLAIAFLATFVLLVEAFVIIYYLRNRNRSKVQLINFIYCSLLVIFST MLFSFALGPYYVSVAYLKY 268
MYPE470 RLLTNKNNPNKIWSFIFFQTILLIFSSICIYILIFKVST-------DIRFEGKSSLEKF----LYSNSKWVILSSIIFFAVCMETLALVFFKKFSKN----FIMCALIISLVCSIS IIFSFVLGPISANEYYKF 241
MGA_0323 RIRKNYASN--FWDFVFQQILIVGFCISGFIFLS-LYNN----------ADTDRFLVR----NNSFGTTLSMVFMALFFVFAELISWLLYLKHRKTKES-IKLFLYVSTMMITLS ILYSLVLGTISINEFLIN 281
MGA_0333 RIRKNYASN--FWDFVFQQILIVGFCISGFIFLS-LYNN----------ADTDRFLVR----NNSFGTTLSMVFMALFFVFAELISWLLYLKHRKTKES-IKLFLYVSTMMITLS ILYSLVLGTISINEFLIN 249
MG313 LVRKNASNY--KVDFWVLQTLLVTFFVVSVTLLLLYLTNN-----EFQQIGTKNFQTADLSVNVVVLQIITLVSFVIFFLVTEGFLGFVYVKKRNKEQ--ALLTIYSLVLMVLMS VVVSILLGTIAAIEFITF 226
MPN448 QVRQSSTNW--KIDFWILQSILIGFFVLTLVLLLMYLTDG-----NFQAAGNQSFGRG-FDVNVHILQGITMAAFISFFVGLEIFLGWKYTKVKKPKE--MILNLYILMMALLMT LVVSLLIGTVASIEYLVF 225
MMOB0810 RNKND--------KLFVIPFSLSIVSLASVFAVFSLGWI----------RNNLKIDGLNGQVTFFAILAISLISVTVIFINIFCFLKYIRTGKN---IYKRIIYILALIIFIK IFARWLWGPIAFIEYYNV 240
MYPU_2960 KNKNL--------KLIFIPITIIVVSLTGGFITFAFYRP----------KKGVDIWIWQEVSFISSIVVSTIVALVSLILIFLFLFLYLKTKKD---IYFRIIIVICLIVSSV VLSRWLMGPYAFTNYLNR 246
CKL_3856 CNKDNKKMRMIVAIFSVLMLLAGICNIAFLAADGVNREFYM----------HVQKENVNTDNMHLISKMLIERTIGTKDPSGNLATYISWVTTAMIGCATLGILLLAADLFLS KKIFHDTSREQIPQLLIAMIGSGLIVTTLNTV 238
lp_0164 T SLTRVLVAN FIVVFILNLFLNSLWLH 124
Cphy_2765 Q SLPNIVGA AVVNQFFCGLVINSYWLS 139
ruler .....190.....200.....210.....220.....230.....240.....250.....260.....270.....280.....290.....300.....310.....320.....330.....340.....350.....360

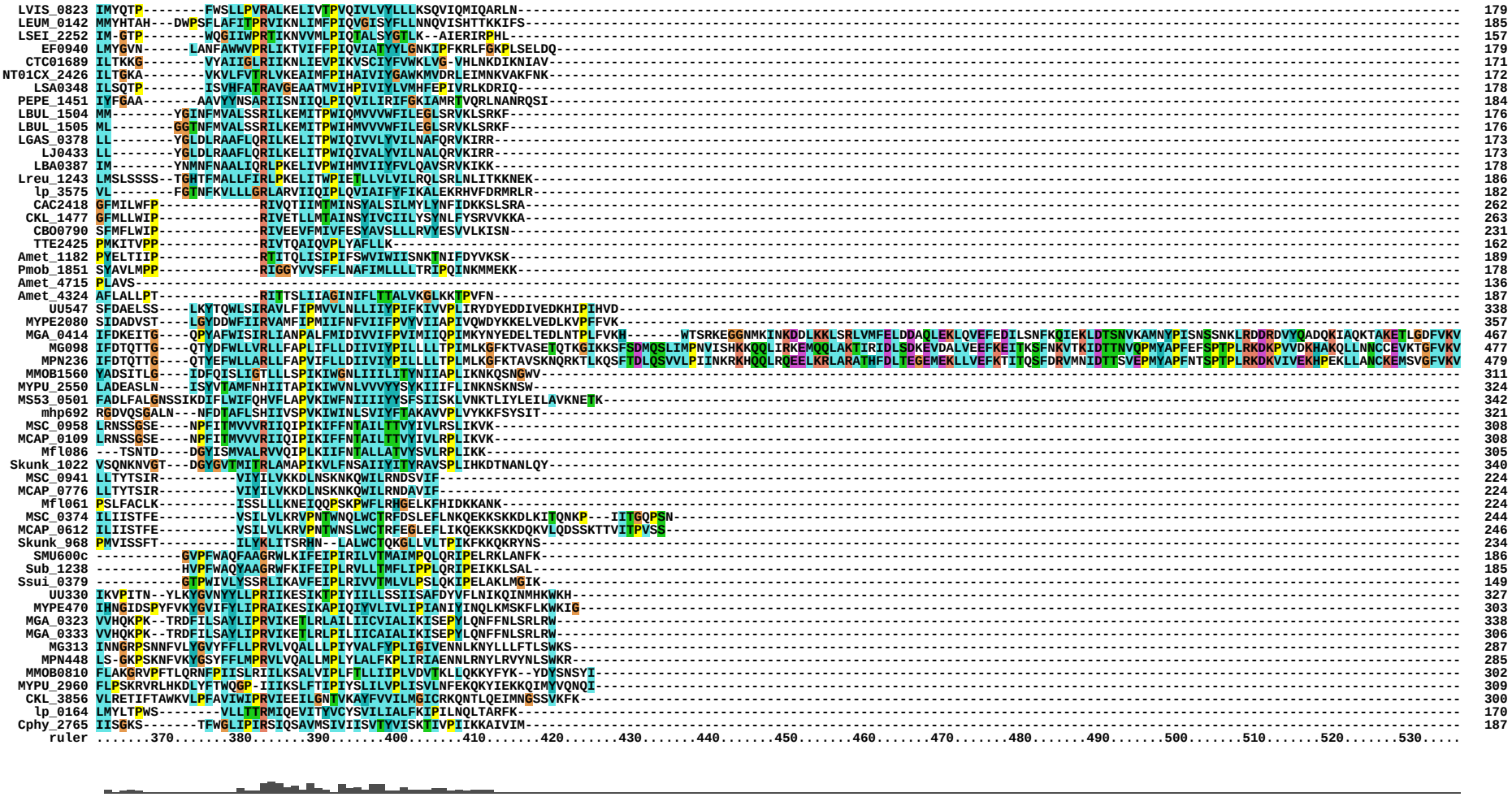
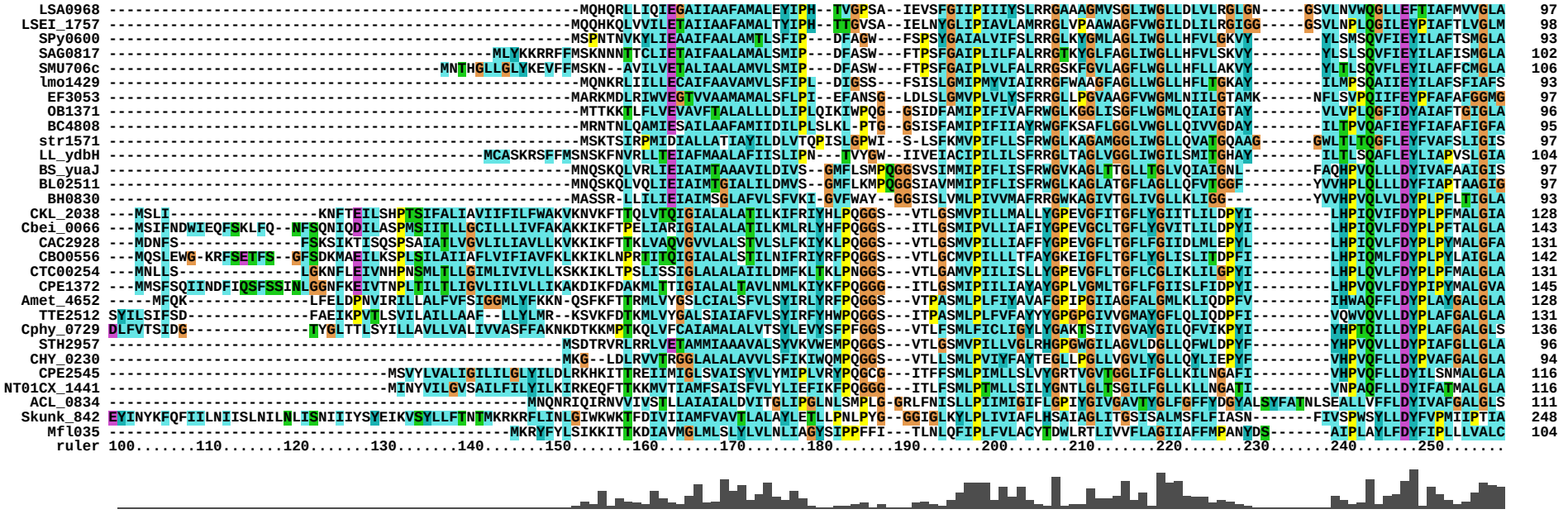
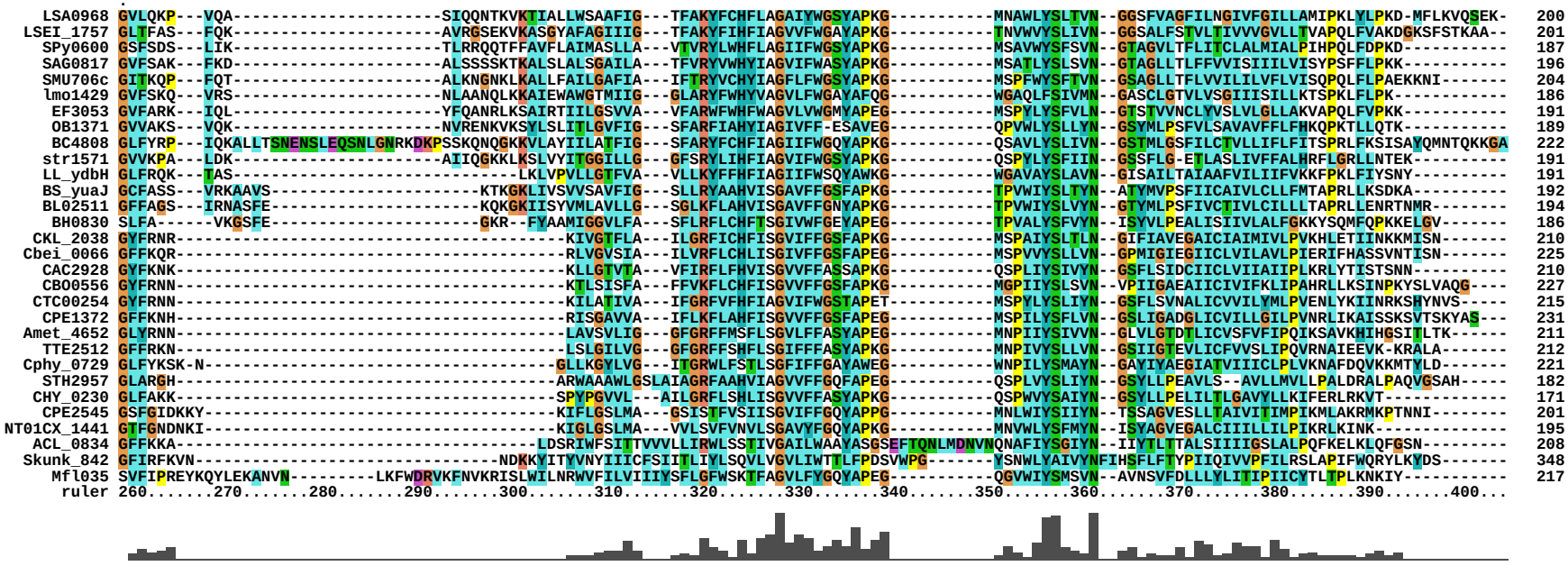

LVIS_0823 IMYQTP--------FWSLLPVRALKELIVTPVQIVLVYLLLKSQVIQMIQARLN---------- 179
LEUM_0142 MMYHTAH---DWPSFLAFITPRVIKNLIMFPIQVGISYFLLNNQVISHTTKKIFS---------- 185
LSEI_2252 IM-GTP--------WQGIIWPRTIKNVVMLPIQTALSYGTLK--AIERIRPHL---------- 157
EF0940 LMYGVN------LANFAWWVPRLIKTVIFFPIQVIATYYLGNKIPFKRLFGKPLSELDQ---------- 179
CTC01689 ILTKKG--------VYAIIGLRIIKNLIEVPIKVSCIYFVWKLVG-VHLNKDIKNIAV---------- 171
NT01CX_2426 ILTGKA--------VKVLFVTRLVKEAIMFPIHAIVIYGAWKMVDRLEIMNKVAKFNK---------- 172
LSA0348 ILSQTP--------ISVHFATRAVGEAATMVIHPIVIYLVMHFEPIVRLKDRIQ---------- 178
PEPE_1451 IYFGAA-------AAVYYNSARIISNIIQLPIQVILIRIFGKIAMRTVQRLNANRQSI---------- 184
LBUL_1504 MM--------YGINFMVALSSRILKEMITPWIQMVVVWFILEGLSRVKLSRKF---------- 176
LBUL_1505 ML--------GGTNFMVALSSRILKEMITPWIHMVVVWFILEGLSRVKLSRKF---------- 176
LGAS_0378 LL--------YGLDLRAAFLQRILKELITPWIQIVVLYVILNAFQRVKIRR---------- 173
LJ0433 LL--------YGLDLRAAFLQRILKELITPWIQIVALYVILNALQRVKIRR---------- 173
LBA0387 IM--------YNMNFNAALIQRLPKELIVPWIHMVIIYFVLQAVSRVKIKK---------- 178
Lreu_1243 LMSLSSSS--TGHTFMALLFIRLPKELITWPIETLLVLVILRQLSRLNLITKKNEK---------- 186
lp_3575 VL--------FGTNFKVLLLGRLARVIIQIPLQVIAIFYFIKALEKRHVFDRMRLR---------- 182
CAC2418 GFMILWFP-------------RIVQTIIMTMINSYALSILMYLYNFIDKKSLSRA---------- 262
CKL_1477 GFMLLWIP-------------RIVETLLMTAINSYIVCIILYSYNLFYSRVVKKA---------- 263
CBO0790 SFMFLWIP-------------RIVEEVFMIVFESYAVSLLLRVYESVVLKISN---------- 231
TTE2425 PMKITVPP-------------RIVTQAIQVPLYAFLLK---------- 162
Amet_1182 PYELTIIP-------------RTITQLISIPIFSWVIWIISNKTNIFDYVKSK---------- 189
Pmob_1851 SYAVLMPP-------------RIGGYVVSFFLNAFIMLLLLTRIPQINKMMEKK---------- 178
Amet_4715 PLAVS---------- 136
Amet_4324 AFLALLPT-------------RITTSLIIAGINIFLTTALVKGLKKTPVFN---------- 187
UU547 SFDAELSS----LKYTQWLSIRAVLFIPMVVLNLLIIYPIFKIVVPLIRYDYEDDIVEDKHIPIHVD---------- 338
MYPE2080 SIDADVST----LGYDDWFIIRVAMFIPMIIFNFVIIFPVYVIIAPIVQWDYKKELVEDLKVPFFVK---------- 357
MGA_0414 IFDKEITG----QPYAFWISIRLIANPALFMIDIVVIFPVIMIIQPIMKYNYEDELTEDLNTPLFVKH-------WTSRKEGGNMKINKDDLKKLSRLVMFELDDAQLEKLQVEFEDILSNFKQIEKLDTSNVKAMNYPISNSSNKLRDDRDVYQADQKIAQKTAKETLGDFVKV 467
MG098 IFDTQTTG----QTYDFWLLVRLLFAPLIFLLDIIVIYPILLLLTPIMLKGFKTVASETQTKGIKKSFSDMQSLIMPNVISHKKQQLIRKEMQQLAKTIRIDLSDKEVDALVEEFKEITKSFNKVTKIDTTNVQPMYAPFEFSPTPLRKDKPVVDKHAKQLLNNCCEVKTGFVKV 477
MPN236 IFDTQTTG----QTYEFWLLARLLFAPVIFLLDIIVIYPILLLLTPLMLKGFKTAVSKNQRKTLKQSFTDLQSVVLPIINKRKHQQLRQEELKRLARATHFDLTEGEMEKLLVEFKTIIQSFDRVMNIDTTSVEPMYAPFNTSPTPLRKDKVIVEKHPEKLLANCKEMSVGFVKV 479
MMOB1560 YADSITLG----IDFQISLIGTLLLSPIKIWGNLIIILITYNIIAPLIKNKQSNGWV---------- 311
MYPU_2550 LADEASLN----ISYVTAMFNHIITAPIKIWVNLVVVYYSYKIIIFLINKNSKNSW---------- 324
MS53_0501 FADLFALGNSSIKDIFLWIFQHVFLAPVKIWFNIIIIYYSFSIISKLVNKTLIYLEILAVKNETIK---------- 342
mhp692 RGDVQSGALN---NFDTAFLSHIIVSPVKIWINLSVIYFTAKAVVPLVYKKFSYSIT---------- 321
MSC_0958 LRNSSGSE----NPFITMVVVRIIQIPIKIFFNTAILTTVYIVLRSLIKVK---------- 308
MCAP_0109 LRNSSGSE----NPFITMVVVRIIQIPIKIFFNTAILTTVYIVLRPLIKVK---------- 308
Mfl086 ---TSNTD----DGYISMVALRVVQIPLKIIFNTALLATVYSVLRPLIKK---------- 305
Skunk_1022 VSQNKNVGT---DGYGVTMITRLAMAPIKVLFNSAIIYITYRAVSPLIHKDTNANLQY---------- 340
MSC_0941 LLTYTSIR----------VIYILVKKDLNSKNKQWILRNDSVIF---------- 224
MCAP_0776 LLTYTSIR----------VIYILVKKDLNSKNKQWILRNDAVIF---------- 224
Mfl061 PSLFACLK----------ISSLLLKNEIQQPSKPWFLRHGELKFHIDKKANK---------- 224
MSC_0374 ILIISTFE----------VSILVLKRVPNTWNQLWCTRFDSLEFLNKQEKKSKKDLKITQNKP---IITGQPSN---------- 244
MCAP_0612 ILIISTFE----------VSILVLKRVPNTWNSLWCTRFEGLEFLIKQEKKSKKDQKVLQDSSKTTVITPVSS---------- 246
Skunk_968 PMVISSFT----------ILYKLITSRHN--LALWCTQKGLLVLTPIKFKKQKRYNS---------- 234
SMU600c ----------GVPFWAQFAAGRWLKIFEIPIRILVTMAIMPQLQRIPELRKLANFK---------- 186
Sub_1238 ----------HVPFWAQYAAGRWFKIFEIPLRVLLTMFLIPPLQRIPEIKKLSAL---------- 185
Ssui_0379 ----------GTPWIVLYSSRLIKAVFEIPLRIVVTMLVLPSLQKIPELAKLMGIK---------- 149
UU330 IKVPITN--YLKYGVNYYLLPRIIKESIKTPIYIILLSSIISAFDYVFLNIKQINMHKWKH---------- 327
MYPE470 IHNGIDSPYFVKYGVIFYLIPRAIKESIKAPIQIYVLIVLIPIANIYINQLKMSKFLKWKIG---------- 303
MGA_0323 VVHQKPK--TRDFILSAYLIPRVIKETLRLAILIICVIALIKISEPYLQNFFNLSRLRW---------- 338
MGA_0333 VVHQKPK--TRDFILSAYLIPRVIKETLRLPILIICAIALIKISEPYLQNFFNLSRLRW---------- 306
MG313 INNGRPSNNFVLYGVYFFLLPRVLVQALLLPIYVALFYPLIGIVENNLKNYLLLFTLSWKS---------- 287
MPN448 LS-GKPSKNFVKYGSYFFLMPRVLVQALLMPLYLALFKPLIRIAENNLRNYLRVYNLSWKR---------- 285
MMOB0810 FLAKGRVPFTLQRNFPIISLRIILKSALVIPLFTLLIIPLVDVTKLLQKKYFYK--YDYSNSYI---------- 302
MYPU_2960 FLPSKRVRLHKDLYFTWQGP-IIIKSLFTIPIYSLILVPLISVLNFEKQKYIEKKQIMYVQNQI---------- 309
CKL_3856 VLRETIFTAWKVLPFAVIWIPRVIEEILGNTVKAYFVVILMGICRKQNTLQEIMNGSSVKFK---------- 300
lp_0164 LMYLTPWS--------VLLTTRMIQEVITYVCYSVILIALFKIPILNQLTARFK---------- 170
Cphy_2765 IISGKS-------TFWGLIPIRSIQSAVMSIVIISVTYVISKTIVPIIKKAIVIM---------- 187
ruler ....370.......380.......390.......400.......410.......420.......430.......440.......450.......460.......470.......480.......490.......500.......510.......520.......530.....

Fig. S3. Sequence alignment of ThiT proteins

Histograms below alignment show the degree of residue conservation at each position




LSA0968 GVLQKP---VQA----------SIQQNTKVKTIALLWSAAFIG---TFAKYFCHFLAGAIYWGSYAPKG----------MNAWLYSLTVN--GGSFVAGFILNGIVFGILLAMIPKLYLPKD-MFLKVQSEK- 200
LSEI_1757 GLTFAS---FQK----------AVRGSEKVKASGYAFAGIIIG---TFAKYFIHFIAGVVFWGAYAPKG----------TNVWVYSLIVN--GGSALFSTVLTIVVVGVLLTVAPQLFVAKDGKSFSTKAA-- 201
SPy0600 GSFSDS---LIK----------TLRRQQTFFAVFLAIMASLLA---VTVRYLWHFLAGIIFWGSYAPKG----------MSAVWYSFSVN--GTAGVLTFLITCLALMIALPIHPQLFDPKD---------- 187
SAG0817 GVFSAK---FKD----------ALSSSSKTKALSLALSGAILA---TFVRYVWHYIAGVIFWASYAPKG----------MSATLYSLSVN--GTAGLLTLFFVVISIIILVISYPSFFLPKK---------- 196
SMU706c GITKQP---FQT----------ALKNGNKLKALLFAILGAFIA---IFTRYVCHYIAGFLFWGSYAPKG----------MSPFWYSFTVN--GSAGLLTFLVVILILVFLVISQPQLFLPAEKKNI------- 204
lmo1429 GVFSKQ---VRS----------NLAANQLKKAIEWAWGTMIIG---GLARYFWHYVAGVLFWGAYAFQG----------WGAQLFSIVMN--GASCLGTVLVSGIIISILLKTSPKLFLPK---------- 186
EF3053 GVFARK---IQL----------YFQANRLKSAIRTIILGSVVA---VFARWFWHFWAGVLVWGMYAPEG----------MSPYLYSFVLN--GTSTVVNCLYVSLVLGLLAKVAPQLFVPKK---------- 191
OB1371 GVVAKS---VQK----------NVRENKVKSYLSLITLGVFIG---SFARFIAHYIAGIVFF-ESAVEG----------QPVWLYSLLYN--GSYMLPSFVLSAVAVFFLFHKQPKTLLQTK---------- 189
BC4808 GLFYRP---IQKALLTSNENSLEQSNLGNRKDKPSSKQNQGKKVLAYIILATFIG---SFARYFCHFIAGIIFWGQYAPKG----------QSAVLYSLIVN--GSTMLGSFILCTVLLIFLFITSPRLFKSISAYQMNTQKKGA 222
str1571 GVVKPA---LDK----------AIIQGKKLKSLVYITGGILLG---GFSRYLIHFIAGVIFWGSYAPKG----------QSPYLYSFIIN--GSSFLG-ETLASLIVFFALHRFLGRLLNTEK---------- 191
LL_ydbH GLFRQK---TAS----------LKLVPVLLGTFVA---VLLKYFFHFIAGIIFWSQYAWKG----------WGAVAYSLAVN--GISAILTAIAAFVILIIFVKKFPKLFIYSNY---------- 191
BS_yuaJ GCFASS---VRKAAVS----------KTKGKLIVSVVSAVFIG---SLLRYAAHVISGAVFFGSFAPKG----------TPVWIYSLTYN--ATYMVPSFIICAIVLCLLFMTAPRLLKSDKA---------- 192
BL02511 GFFAGS---IRNASFE----------KQKGKIISYVMLAVLLG---SGLKFLAHVISGAVFFGNYAPKG----------TPVWIYSLVYN--GTYMLPSFIVCTIVLCILLLTAPRLLENRTNMR-------- 194
BH0830 SLFA------VKGSFE----------GKR--FYAAMIGGVLFA---SFLRFLCHFTSGIVWFGEYAPEG----------TPVALYSFVYN--ISYVLPEALISIIVLALFGKKYSQMFQPKKELGV------- 186
CKL_2038 GYFRNR----------KIVGTFLA---ILGRFICHFISGVIFFGSFAPKG----------MSPAIYSLTLN--GIFIAVEGAICIAIMIVLPVKHLETIINKKMISN------- 210
Cbei_0066 GFFKQR----------RLVGVSIA---ILVRFLCHLISGIVFFGSFAPEG----------MSPVVYSLLVN--GPMIGIEGIICLVILAVLPIERIFHASSVNTISN------- 225
CAC2928 GYFKNK----------KLLGTVTA---VFIRFLFHVISGVVFFASSAPKG----------QSPLIYSIVYN--GSFLSIDCIICLVIIAIIPLKRLYTISTSNN---------- 210
CBO0556 GYFRNN----------KTLSISFA---FFVKFLCHFISGVVFFGSFAPKG----------MGPIIYSLSVN--VPIIGAEAIICIVIFKLIPAHRLLKSINPKYSLVAQG---- 227
CTC00254 GYFRNN----------KILATIVA---IFGRFVFHFIAGVIFWGSTAPET----------MSPYLYSLIYN--GSFLSVNALICVVILYMLPVENLYKIINRKSHYNVS----- 215
CPE1372 GFFKNH----------RISGAVVA---IFLKFLAHFISGVVFFGSFAPEG----------MSPILYSFLVN--GSLIGADGLICVILLGILPVNRLIKAISSKSVTSKYAS--- 231
Amet_4652 GLYRNN----------LAVSVLIG---GFGRFFMSFLSGVLFFASYAPEG----------MNPIIYSIVVN--GLVLGTDTLICVSFVFIPQIKSAVKHIHGSITLTK------ 211
TTE2512 GFFRKN----------LSLGILVG---GFGRFFSHFLSGIFFFASYAPKG----------MNPIVYSLLVN--GSIIGTEVLICFVVSLIPQVRNAIEEVK-KRALA------- 212
Cphy_0729 GLFYKSK-N----------GLLKGYLVG---ITGRWLFSTLSGFIFFGAYAWEG----------WNPILYSMAYN--GAYIYAEGIATVIIICLPLVKNAFDQVKKMTYLD------- 221
STH2957 GLARGH----------ARWAAAWLGSLAIAGRFAAHVIAGVVFFGQFAPEG----------QSPLVYSLIYN--GSYLLPEAVLS--AVLLMVLLPALDRALPAQVGSAH----- 182
CHY_0230 GLFAKK----------SPYPGVVL---AILGRFLSHLISGVVFFASYAPKG----------QSPWVYSAIYN--GSYLLPELILTLGAVYLLKIFERLRKVT---------- 171
CPE2545 GSFGIDKKY----------KIFLGSLMA---GSISTFVSIISGVIFFGQYAPPG----------MNLWIYSIIYN--TSSAGVESLLTAIVITIMPIKMLAKRMKPTNNI-------- 201
NT01CX_1441 GTFGNDNKI----------KIGLGSLMA---VVLSVFVNVLSGAVYFGQYAPKG----------MNVWLYSFMYN--ISYAGVEGALCIIILLILPIKRLKINK---------- 195
ACL_0834 GFFKKA----------LDSRIWFSITTVVVLLIRWLSSTIVGAILWAAYASGSEFTQNLMDNVNQNAFIYSGIYN--IIYTLTTALSIIIGSLALPQFKELKLQFGSN--------- 208
Skunk_842 GFIRFKVN----------NDKKYITYVNYIIICFSIITLIYLSQVLVGVLIWTTLFPDSVWPG-------YSNWLYAIVYNFIHSFLFTYPIIQIVVPFILRSLAPIFWQRYLKYDS------- 348
Mfl035 SVFIPREYKQYLEKANVN---------LKFWDRVKFNVKRISLWILNRWVFILVIIIYSFLGFWSKTFAGVLFYGQYAPEG----------QGVWIYSMSVN--AVNSVFDLLLYLITIPIICYTLTPLKNKIY---------- 217
ruler 260.....270.....280.....290.....300.....310.....320.....330.....340.....350.....360.....370.....380.....390.....400...

Supplemental Table 1

Locus tags of FolT and ThiT proteins identified in various bacterial genomes

FolT proteins

Organism	Locus tag
Clostridium acetobutylicum ATCC 824	CAC2418
Clostridium botulinum A str. ATCC 3502	CBO0790
Clostridium kluyveri DSM 555	CKL 3856
Clostridium kluyveri DSM 555	CKL 1477
Clostridium novyi NT	NT01CX_2426
Clostridium phytofermentans ISDg	Cphy_2765
Clostridium tetani E88	CTC01689
Enterococcus faecium DO	EFAE 0455
Enterococcus faecalis V583	EF0940
Lactobacillus acidophilus NCFM	LBA0387
Lactobacillus brevis ATCC 367	LVIS_0823
Lactobacillus casei ATCC 334	LSEI 2252
Lactobacillus delbrueckii subsp. *bulgaricus* ATCC BAA-365	LBUL 1505
Lactobacillus delbrueckii subsp. *bulgaricus* ATCC BAA-365	LBUL 1504
Lactobacillus gasseri ATCC 33323	LGAS 0378
Lactobacillus johnsonii NCC 533	LJ0433
Lactobacillus plantarum WCFS1	lp_3575
Lactobacillus plantarum WCFS1	lp_0164
Lactobacillus reuteri F275	LREU 1243
Lactobacillus sakei subsp. sakei 23K	LSA0348
Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293	LEUM_0142
Mesoplasma florum L1	MFL061
Mesoplasma florum L1	Mfl086
Mycoplasma capricolum subsp. capricolum ATCC 27343	MCAP_0612
Mycoplasma capricolum subsp. capricolum ATCC 27343	MCAP_0776
Mycoplasma capricolum subsp. capricolum ATCC 27343	MCAP_0109
Mycoplasma gallisepticum R	MGA 0333
Mycoplasma gallisepticum R	MGA 0323
Mycoplasma gallisepticum R	MGA 0414
Mycoplasma genitalium G37	MG_313
Mycoplasma genitalium G37	MG_098
Mycoplasma hyopneumoniae 232	MHP692
Mycoplasma mobile 163K	MMOB0810
Mycoplasma mobile 163K	MMOB1560
Mycoplasma mycoides subsp. mycoides SC str. PG1	MSC 0374
Mycoplasma mycoides subsp. mycoides SC str. PG1	MSC 0941
Mycoplasma mycoides subsp. mycoides SC str. PG1	MSC 0958
Mycoplasma penetrans HF2	MYPE470
Mycoplasma penetrans HF2	MYPE2080
Mycoplasma pneumoniae M129	MPN448
Mycoplasma pneumoniae M129	MPN236
Mycoplasma pulmonis UAB CTIP	MYPU 2960
Mycoplasma pulmonis UAB CTIP	MYPU 2550
Mycoplasma synoviae 53	MS53 0501
Petrotoga mobilis SJ95	PMOB 1851
Pediococcus pentosaceus ATCC 25745	PEPE 1451
Spiroplasma kunkelii CR2-3X	Skunk_968
Spiroplasma kunkelii CR2-3X	Skunk_1022
Streptococcus mutans UA159	SMU.600c
Streptococcus suis 89/1591	Ssui_0379
Streptococcus uberis 0140J	SUB1238
Ureaplasma parvum serovar 3 ATCC 700970	UU330
Ureaplasma parvum serovar 3 ATCC 700970	UU547

ThiT proteins

Organism	Locus tag
Acholeplasma laidlawii PG-8A	ACL_0834
Alkaliphilus metalliredigens QYMF	Amet_4652
Bacillus halodurans C-125	BH0830
Bacillus cereus ATCC 14579	BC4808
Bacillus licheniformis ATCC 14580	BL02511
Bacillus subtilis subsp. subtilis str. 168	BSU30990
Carboxydothermus hydrogenoformans Z-2901	CHY_0230
Clostridium acetobutylicum ATCC 824	CAC2928
Clostridium beijerinckii NCIMB 8052	Cbei_0066
Clostridium botulinum A str. ATCC 3502	CBO0556
Clostridium kluyveri DSM 555	CKL_2038
Clostridium perfringens str. 13	CPE2545
Clostridium perfringens str. 13	CPE1372
Clostridium phytofermentans ISDg	Cphy_0729
Clostridium tetani E88	CTC00254
Enterococcus faecalis V583	EF3053
Lactobacillus casei ATCC 334	LSEI_1757
Lactococcus lactis subsp. lactis Il1403	L116212
Listeria monocytogenes EGD-e	lmo1429
Oceanobacillus iheyensis HTE831	OB1371
Spiroplasma kunkelii CR2-3X	Skunk_842
Streptococcus agalactiae 2603V/R	SAG0817
Streptococcus mutans UA159	SMU.706c
Streptococcus pyogenes M1 GAS	YP_602012
Streptococcus thermophilus CNRZ1066	str1571
Symbiobacterium thermophilum IAM 14863	STH2957
Thermoanaerobacter tengcongensis MB4	TTE2512